View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by AIS Electronic Library (AISeL)

Association for Information Systems
AIS Electronic Library (AISeL)

CONF-IRM 2012 Proceedings

International Conference on Information Resources Management (CONF-IRM)

5-2012

Framework for connections on Facebook: An exploratory study

František Sudzina
Aarhus University, sudzina@business.aau.dk

Follow this and additional works at: http://aisel.aisnet.org/confirm2012

Recommended Citation
Sudzina, František, "Framework for connections on Facebook: An exploratory study" (2012). *CONF-IRM 2012 Proceedings*. 55.
http://aisel.aisnet.org/confirm2012/55

This material is brought to you by the International Conference on Information Resources Management (CONF-IRM) at AIS Electronic Library (AISeL). It has been accepted for inclusion in CONF-IRM 2012 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

Framework for connections on Facebook: An exploratory study

František Sudzina
Aarhus University
fransu@asb.dk

Abstract

There is a substantial amount of current information systems and marketing research focused on social networking sites, most frequently on Facebook. Often, these studies utilize available metadata on user on-line behavior, such as what links the users clicked on. In order to better understand behavior of Facebook users, it makes sense to investigate also whom the users connect to. It is possible to hypothesize that behavior of people, who connect only to relatives on Facebook, differs from behavior of people, who are connected only to their classmates. The paper offers a framework of Facebook connections and a proposal how to group these connections.

Keywords

Facebook, Connection Types, Empirical Research.

1. Introduction

The last decade may be characterized by moving focus from Web to Web 2.0. Wikis, blogs, and social networking sites are used now by most of the Internet users. While access to wikis and blogs is rather open, the access to the content of social networking sites is usually restricted to a small set of people handpicked by social networking sites users. In other words, social networking sites are more personal. This attracts interest of researchers; the most researched social networking site so far is Facebook. The usually investigated topic is the user behavior; sometimes only with a goal to understand it (e.g. what kind of messages and pictures users post on walls), other times with a goal to influence it as well.

An example for the latter can be the problem posed by Domingos and Richardson (2001) - if we can try to convince a subset of individuals to adopt a new product or innovation, and the goal is to trigger a large cascade of further adoptions, which set of individuals should we target? It is possible to suggest a solution based on strength of particular connections, e.g. like it was proposed by Korfiatis and Sicilia (2007). It, certainly, describes an important property of the connection but it lacks a qualitative information on the relationship between the two users. Although one could infer e.g. from the communications (chat and mails) how strong is the relationship, maybe relatives (even with very little communication) influence a purchase decision more than people, with whom the user communicates on a frequent basis.

Going back to the previous example, a user that connects only to relatives, will probably predominantly post messages about and pictures of himself/herself and of his/her children and/or parents. While a user, who is connected to his/her colleagues from work will probably post messages and pictures relevant to his/her company, and/or his professional interests. So, it is

possible to propose that research of Facebook user behavior should include several connection types of as an independent variable.
This paper offers such a framework with nine connection types. (The term "connection" will be used throughout the paper for what Facebook terms "friend" because "friend" is one of the connection types proposed in this framework.) Since adding nine additional variables to a model increases the degree of freedom by nine, such a step would most likely to lead to worse p-values given a constant sample size. Therefore, the aim of the paper is also to propose how to cluster these connection types into meaningful groups. The next section describes the data and the methodology used for clustering. The third section then discusses the proposed groups of connection types, and the final section provides conclusions.

2. Data and methodology

A questionnaire was used to collect data on Facebook connection types. The survey was conducted in February 2011. Respondents (154 in total) were students of the School of Business and Social Sciences, Aarhus University (57), of Copenhagen Business School (41), and of the Faculty of Business Economics, the University of Economics Bratislava (56). Two students from the latter school stated that they did not have a Facebook account. All the others stated that they had a Facebook account. So, the effective sample size is 152. With regards to gender, 53 respondents were male, 98 respondents were female and one respondent did not specify the gender. The proportion of the males and females in the sample is consistent with the gender proportion at business schools both in Denmark and in Slovakia. With regards to the national origin, 32 students were from Denmark, 54 from Slovakia, 45 from the rest of Europe, and 21 from outside Europe.
The question used to collect data on connection types was "Whom do you connect to (send and accept invitations to/from) on Facebook? (Please tick all that that apply.)" The possibilities were (the wording in Tables 2 and 3 will be shortened but it will be provided in the same order):

- Classmates from primary school;
- Classmates from secondary/high school;
- Classmates from college and/or university;
- Acquaintances, who lived or live close to your place;
- Acquaintances, whom you met during your exchange program, at a summer course, a training, a language course;
- Colleagues from office;
- Friends;
- Relatives; and
- Others.

Although the framework may seem to be skewed towards university students, it should not be considered as a weakness. First of all, researchers often opt for convenience samples, and the university students are then often the first choice. So, it is possible that researchers will look for a framework that suits university students. Secondly, the framework suits well also university graduates. It is possible that people, who graduated prior to 2007 or shortly after (when the Facebook became more widespread) are less likely to connect to acquaintances, whom they met during an exchange program or at a summer course. It may be assumed because they would have weaker ties to these people, they would not stay in touch with them afterwards, thus they would

not connect to them once they joined Facebook. But even people, who graduated a longer time ago, may still have attended a training or a language course, so this connection type would still be relevant for them. With regards to people, who have not attended the university, the third connection type - classmates from college and/or university – would be irrelevant. But all the remaining eight connection types should be still relevant for them.

Answers were coded as 0 – no tick (the respondent does not accept nor requests connection to the particular group of people), and 1 – tick (the respondent accepts and requests connection to the particular group of people). Pearson correlations coefficients between connection types range from -0.066 to 0.443. Overall, the connection types are somewhat correlated, the Cronbach's alpha is 0,619 (the standardized Cronbach's alpha is 0,667). Clustering using Krajci's (2003) fuzzy concepts ended up with rather many clusters. This can be also interpreted as a result of the correlation between connection types. PCA (principal component analysis), defined by Pearson (1901), using Varimax rotation, proposed by Kaiser (1958), will be used to group the connection types. Correlations in data hint that the result of the PCA will not be ideal. Despite this expectation, the preferred outcome is the set of factors that are orthogonal to each other. The main goal is to create factors that cover all connection types.

3. Results and discussion

It is appropriate to use PCA for the data because it is possible to reject the hypothesis that all eigenvalues are equal (Phi=0,216734, Log(Det|R|) = -1.297, Bartlett test=190.83, DF=36, and the related p-value < 0.001). First, let us look into eigenvalues after Varimas rotation in Table 1.

No.	Eigenvalue	Individual Percent	Cumulative Percent
1	2.211	24.57	24.57
2	1.392	15.47	40.04
3	1.291	14.34	54.38
4	0.983	10.93	65.30
5	0.920	10.23	75.53
6	0.670	7.44	82.97
7	0.585	6.50	89.47
8	0.513	5.70	95.17
9	0.435	4.83	100.00

Table 1: Eigenvalues after Varimax Rotation

There are three eigenvalues higher than 1 in Table 1. This implies that there should be three factors that would together explain 54.38 of variance. These factor loadings are provided in Table 2. (Loadings higher than 0.5 are in bold.)

The first factor includes all kinds of classmates and relatives, the second factor includes colleagues and other connections, the third factor includes acquaintances, whom you met during your exchange program, at a summer course, a training, a language course and friends. But neighbors are not correlated enough with any of the three factors. Since the main goal of the analysis was to group *all* the connection types, we may simply introduce the fourth factor that would include neighbors or we may rerun the PCA with a lower threshold for the eigenvalue (let

us say 0.98 because of the fourth eigenvalue in Table 1 is 0.983). Factor loadings for the latter setting are provided in Table 3.

Variables	Factor1	Factor2	Factor3
Primary school classmates	**0.770**	0.227	-0.289
Secondary school classmates	**0.642**	0.348	0.053
University classmates	**0.655**	-0.029	0.422
Neighbors	0.348	0.110	0.092
Short stay connections	-0.098	0.441	**0.708**
Office colleagues	0.167	**0.752**	0.049
Friends	0.305	-0.159	**0.698**
Relatives	**0.717**	-0.048	0.161
Others	0.105	**0.648**	-0.005

Table 2: Factor Loadings after Varimax Rotation (Minimum Eigenvalue = 1)

Variables	Factor1	Factor2	Factor3	Factor4
Primary school classmates	**0.618**	0.309	-0.391	0.317
Secondary school classmates	**0.553**	0.389	-0.021	0.284
University classmates	**0.723**	0.016	0.288	0.099
Neighbors	0.091	-0.013	0.094	**0.935**
Short stay connections	-0.051	0.337	**0.766**	0.177
Office colleagues	0.156	**0.798**	0.080	-0.135
Friends	0.478	-0.160	**0.603**	-0.063
Relatives	**0.787**	0.057	0.008	-0.068
Others	0.022	**0.647**	0.045	0.102

Table 3: Factor Loadings after Varimax Rotation (Minimum Eigenvalue = 0.98)

The first three factors in Table 3 are the same as in Table 2. The fourth factor contains neighbors. So, the suggested grouping of connection types is:

- People, with whom one spent a lot of time
    - Classmates from primary school
    - Classmates from secondary/high school
    - Classmates from college and/or university
    - Relatives
- Less personal connections
    - Colleagues from office
    - Others
- More personal connections
    - Acquaintances, whom you met during your exchange program, at a summer course, a training, a language course
    - Friends
- Neighbors
    - Acquaintances, who lived or live close to your place

The four factors, although grouped using statistics, actually make sense. The first factor includes people, whom one knows for a long time and knows them well but this does not imply that one enjoys spending time with them. The second factor contains people, whom one does not necessarily know well nor for a long time and does not necessarily enjoy spending time with them. The third factor includes people, whom one knows well (unclear for how long) and probably enjoys spending time with them. The fourth factor contains people, whom one probably knows for a long time but not well.

4. Conclusions

The paper proposed a framework for connections, which should be applicable not only to Facebook but at least to some (if not all) social networking sites. Based on the collected data, it is possible to suggest clustering the connection types into four groups - people, with whom one spent a lot of time (classmates from primary and secondary school and college/university and relatives), less personal connections (colleagues from office and others), more personal connections (acquaintances, whom one met during an exchange program, at a summer course, a training, a language course and friends), and neighbors. These groups can be used for structural equations modeling or partial least square regressions.

References

Domingos, P. and Richardson, M. (2001) "Mining the Network Value of Customers" in *Knowledge Discovery and Data Mining*, pp. 57-66.

Kaiser, H.F. (1958) "The varimax criterion for analytic rotation in factor analysis", *Psychometrika*, 23(3), pp. 187-200.

Korfiatis, N. and Sicilia, M.A. (2007) Title: "Social Measures and Flexible Navigation on Online Contact Networks" in *IEEE International Conference on Fuzzy Systems*, London, UK: Imperial College, pp. 1366-1371.

Krajci, S. (2003) "Cluster based efficient generation of fuzzy concepts", *Neural Network World*, 13(5), pp. 521-530.

Pearson, K. (1901) "On Lines and Planes of Closest Fit to Systems of Points in Space", *Philosophical Magazine*, 2(6), pp. 559-572.